

Research Article

Target Detection Using Nonsingular Approximations for a Singular Covariance Matrix

Nir Gorelik,¹ Dan Blumberg,² Stanley R. Rotman,¹ and Dirk Borghys³

¹Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

²Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

³Signal and Image Centre, Royal Military Academy, 1000 Brussels, Belgium

Correspondence should be addressed to Nir Gorelik, nir.gorelik@gmail.com

Received 1 April 2012; Accepted 7 June 2012

Academic Editor: Xiaofei Hu

Copyright © 2012 Nir Gorelik et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate covariance matrix estimation for high-dimensional data can be a difficult problem. A good approximation of the covariance matrix needs in most cases a prohibitively large number of pixels, that is, pixels from a stationary section of the image whose number is greater than several times the number of bands. Estimating the covariance matrix with a number of pixels that is on the order of the number of bands or less will cause not only a bad estimation of the covariance matrix but also a singular covariance matrix which cannot be inverted. In this paper we will investigate two methods to give a sufficient approximation for the covariance matrix while only using a small number of neighboring pixels. The first is the quasilocal covariance matrix (QLRX) that uses the variance of the global covariance instead of the variances that are too small and cause a singular covariance. The second method is sparse matrix transform (SMT) that performs a set of K-givens rotations to estimate the covariance matrix. We will compare results from target acquisition that are based on both of these methods. An improvement for the SMT algorithm is suggested.

1. Introduction

The most widely used algorithms for target detection are traditionally based on the covariance matrix [1]. This matrix estimates the direction and magnitude of the noise in an image. In the equation for a matched filter presented in [1] we have

$$R = t^T \Phi_G^{-1} (x - m), \quad (1)$$

x is the examined pixel, m is the estimate of that pixel based on the surroundings, Φ_G is the global covariance matrix, and t is the target signature. In words, we can say that our matched filter for target detection will detect the target in a particular pixel x if x is different than its surroundings ($x - m$), unlike the noise (controlled by Φ_G^{-1}) and in the direction of the target. If the target signature is unknown, then the RX algorithm uses the target residual ($x - m$) as its own match, that is,

$$R = (x - m)^T \Phi_G^{-1} (x - m), \quad (2)$$

Φ_G is traditionally calculated as follows:

$$\Phi_G = \frac{1}{N} \sum_{i=1}^N (x_i - m)(x_i - m)^T. \quad (3)$$

Although the equation is theoretically justified if the background is stationary, it is often used in cases where this is not true.

In target detection, the image is not normally statistically stationary; it will however have quasistationary “patches” which connect to each other at the edges. When one estimates the mean and covariance matrix of the background of a particular pixel, the local neighboring pixels will have provided a better estimate than the pixels of the entire image. In [2], we show that much better results can be obtained if one uses a “quasilocal covariance matrix” (QLRX). In general terms, it uses the eigenvectors of the overall global matrix, but the eigenvalues are taken locally. This tends to lower the matched filter scores at edges in the data (when the image is going from one stationary distribution to another), but allows for accurate detection in less noisy areas.

TABLE 1: Datasets information—OBP1 and OBP2 are parts of OBP.

Name	Site	Sensor name	No. bands	Waveband (μm)	Spat. Res. (m)	Scene description
OBP	Oberpfaffenhofen (Ge)	Hymap	126	0.44–2.45	4	Airfield with agricultural area around
OBP1	Oberpfaffenhofen (Ge)	Hymap	126	0.44–2.45	4	Agricultural area
OBP2	Oberpfaffenhofen (Ge)	Hymap	126	0.44–2.45	4	Agricultural area

Mathematically, the rotation matrix T will be all the selected rotations combined:

$$T = \prod_{k=K-1}^0 E_k = E_{k-1}E_{k-2} \cdots E_0, \quad \text{when } K < \binom{N}{2}. \quad (12)$$

The variances will be the variances of the local covariance matrix in the direction of the rotation matrix T , $\Lambda_{\text{SMT}} = \text{diag}(T^T \Phi_L T)$ and the inversed covariance matrix will be $\Phi_{\text{SMT}}^{-1} = T \Lambda_{\text{SMT}}^{-1} = T^T$. to decide what rotation matrix is best we use the maximum likelihood covariance estimation and “leave-third-out” cross-validation. (Note that the use of leave-one-out cross-validation will give better results but will cost much more in computational efforts).

We divide the group of pixels into three groups. We take one third to be the tested pixels $Y \in R^{(n/3) \times p}$, and we use the other two thirds to make the approximation of the covariance Φ_{SMT} . After every rotation we calculate the likelihood of covariance Φ_{SMT} to describe correctly the group Y . We have processed to data to make sure that Y is zero mean as follows:

$$P_{\Phi_{\text{SMT}}}(Y) = \frac{1}{(2\pi)^{p/2} |\Phi_{\text{SMT}}|^{1/2}} \exp\left\{-\frac{1}{2} \text{tr}\{Y^T \Phi_{\text{SMT}}^{-1} Y\}\right\}. \quad (13)$$

We do this three times, each time another third is being taken out as the test data Y ; after combining the results of the three tests, we find the number of rotations that gives the best result (based on the highest value of $P(Y)$); we then use the full set and this number of rotations to get the final approximation of the covariance matrix (see Figure 1).

To select every time the rotations that will make the biggest improvement, we perform greedy minimization, that is, always choosing the next rotation that will contribute most to reduce the correlation between data along the axis of the matrix as follows:

$$(i_k, j_k) \leftarrow \arg \max_{i,j} \frac{S_{ij}^2}{S_{ii}S_{jj}}, \quad (14)$$

S is the current covariance matrix, (i, j) are indices of two rows in the matrix, and S_{ij}, S_{ii}, S_{jj} are the members in the matrix with those indices.

After we calculate the covariance matrix for SMT, we can use it for anomaly detection:

$$\text{RX}_{\text{SMT}} = (x - m_L)^T \Phi_{\text{SMT}}^{-1} (x - m_L), \quad (15)$$

x —the tested pixel.

m_L —the mean of the selected shrouding pixels in PCA subspace.

For subpixel targets as we stated previously, it will be better to use the mean of 8 neighbors m_8 .

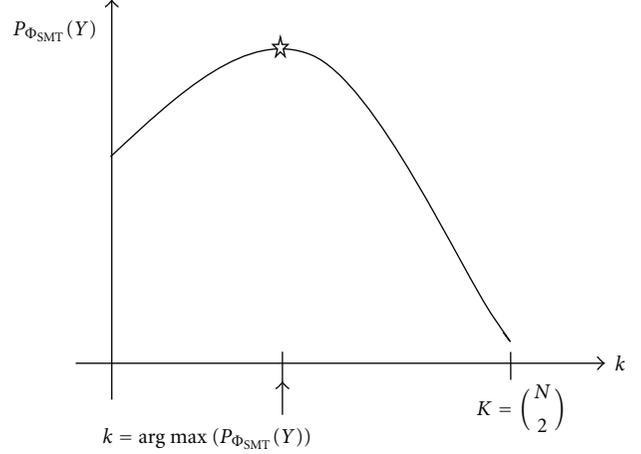


FIGURE 1: The probability that Φ_{SMT} describes Y correctly after k rotations; the k that will be chosen is the one that gives the maximum probability.

3. Dataset

Two datasets were used (Figure 2); a description of their origin can be found in Table 1 and in greater detail in [4].

The two data cubes (OBP1 and OBP2) are real data from the Hymap sensor in which anomalies were inserted artificially by linearly mixing the spectra of a green paint pixel with the original background pixel. For display purposes, in Figure 2, images with full-pixel paint spectra are shown. For the evaluation of anomaly detection results, images with a mixing ratio of 0.33 ($P = 33\%$) were used.

4. Results

We now wish to compare the SMT and QLRX algorithms. We will perform RX anomaly detection (2) using the covariance matrices given by each of the algorithms.

Since the dataset being used contains implanted subpixel targets without any danger of overlap into neighboring pixels, the mean in the calculation of (2) was always the mean of the eight nearest neighbors. However, we must consider the correct neighborhood for the calculation of the covariance matrix for SMT that provides the best results.

The first test was done using only the nearest 8 neighbors for the approximation of the covariance; in this test, it is very easy to see that QLRX results are superior to the SMT results. The ROC curves are given in Figure 3. We assume that the area of the target (and of any examined pixel) consists of the square region of dimension OWS by OWS (outer window). The target area itself has area GWS by GWS (guard window);

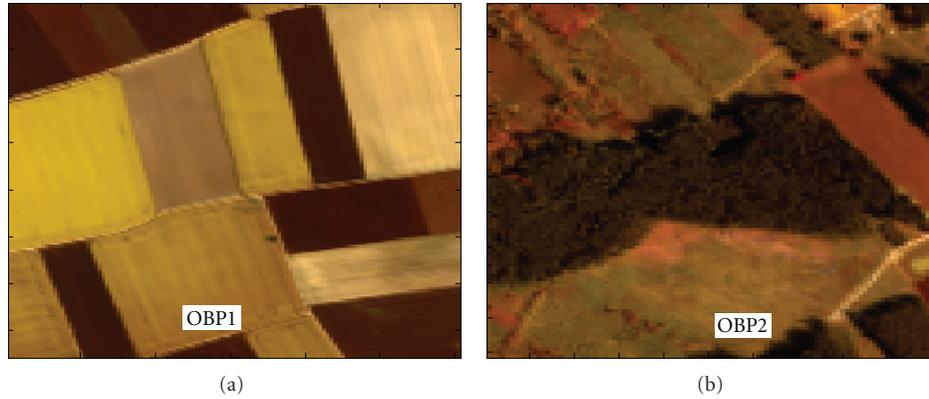


FIGURE 2: RGB composite of the original data cubes. From left to right: OBP1, OBP2.

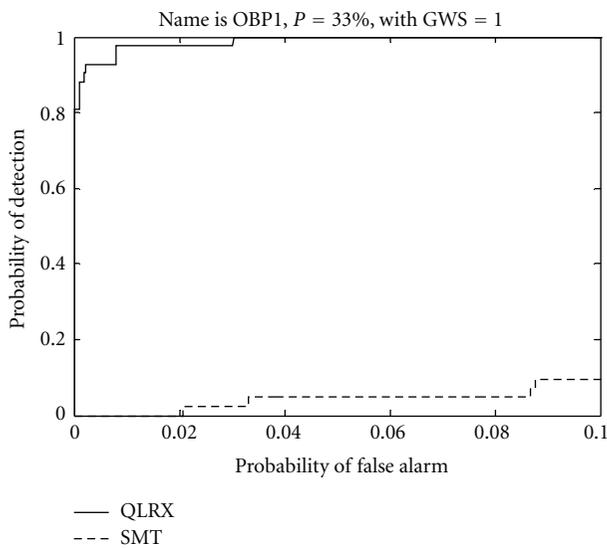


FIGURE 3: Results of RX algorithm using QLRX and SMT on dataset OBP1 with the stated OWS = 3 and GWS = 1.

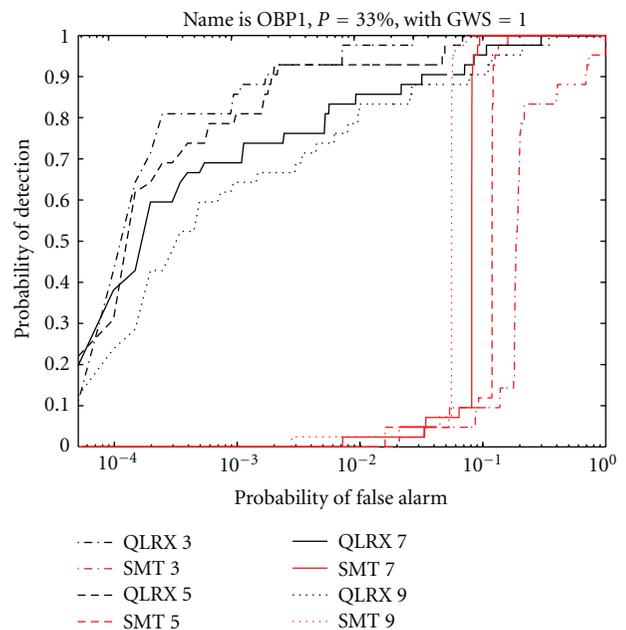


FIGURE 4: OBP1 results with OWS given by the number in the legend.

for subpixel targets GWS will equal 1. Thus the neighboring pixels are those pixels which are located in the square set of pixels in the area OWS by OWS not contained in the inner GWS by GWS matrix.

In this picture is the result for $P = 33\%$, but the tests for 10, 25, 50, and 100 percents gave similar results.

Results from the OBP2 dataset were comparable.

When the OWS is larger, the results of the SMT improve dramatically.

For the dataset OBP1 we can see a large improvement as OWS increases. For this dataset QLRX gives better results, especially in the low CFAR (constant false alarm rate).

For the dataset OBP2, the differences between QLRX and SMT are reduced but still QLRX performs better in the low CFAR regime (see Figure 5).

Similar results were received for the cases in which 10, 25, 50, or 100 percent of the target is in the pixel.

The SMT has two main difficulties: first, the algorithm calculates a new covariance matrix at every point. This calculation needs a sequential set of rotations based on the

training set followed by evaluations of the test set. Both sets are taken from the pixel surrounding only, so none of the information outside the selected group is used in the calculation. In QLRX, the eigenvectors are the same for all points (the eigenvectors of the global covariance). All that we need to do is measure the variance in the local area in the spectral direction of the eigenvalues and calculate the new covariance matrix. Second, the calculation of the SMT itself is highly dependent on the size of the “local” area. While a larger area improves the results, it also increases the time for calculation (see Table 2).

5. Improvements for SMT

A small change in the published method for doing SMT could lead to a large improvement.

TABLE 2: This table shows the time it took to complete the calculation of a dataset.

“Name”	“OWS”	“GWS”	“QLRX” time in seconds	“SMT” time in seconds	Time ratio
“OP1.T1.S10”	9	1	26	3845	148
“OP1.T1.S33”	9	1	26	3794	146
“OP1.T1.S100”	9	1	25	3820	153
“OP1.T1.S10”	7	1	25	2781	111
“OP1.T1.S33”	7	1	24	2769	115
“OP1.T1.S100”	7	1	24	2770	115
“OP1.T1.S10”	5	1	23	1965	85
“OP1.T1.S33”	5	1	23	1969	86
“OP1.T1.S100”	5	1	23	1965	85
“OP1.T1.S10”	3	1	23	1195	52
“OP1.T1.S33”	3	1	22	1187	54
“OP1.T1.S100”	3	1	23	1201	52
“OP2.T1.S10”	9	1	14	2135	153
“OP2.T1.S33”	9	1	14	2105	150
“OP2.T1.S100”	9	1	14	2095	150
“OP2.T1.S10”	7	1	14	1480	106
“OP2.T1.S33”	7	1	14	1481	106
“OP2.T1.S100”	7	1	14	1464	105
“OP2.T1.S10”	5	1	13	961	74
“OP2.T1.S33”	5	1	14	962	69
“OP2.T1.S100”	5	1	13	960	74
“OP2.T1.S10”	3	1	12	556	46
“OP2.T1.S33”	3	1	13	555	43
“OP2.T1.S100”	3	1	14	552	39

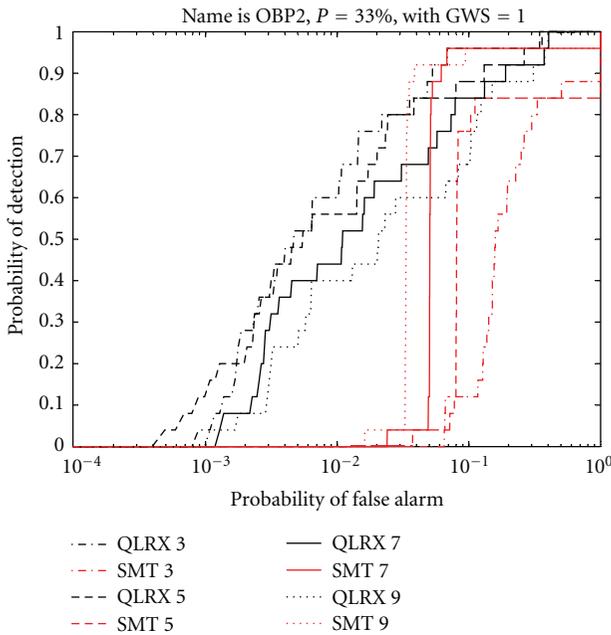


FIGURE 5: Similar to Figure 4 for the OBP2 dataset.

In the original algorithm, the initial assumed axes of the covariance matrix are in the direction of the original dataset.

Then the axes are rotated in pairs into the directions of the “local eigenvectors” to create new covariance matrices.

When we stop, some of the axes will be almost the same direction as the local covariance eigenvectors and some will be closer to the direction of the original axes.

Now since the original directions were random, that is, not related to the correlations between the axes, it is easy to see that there is no reason that this should be optimum. In particular, would it not make more sense to start from the global eigenvectors and rotate into the local ones? Another benefit we will get from this approach is that we will start the rotation from a condition that most probably will be closer to the optimum point (see Figure 6); within fewer rotations, we will get to the maximum likelihood. We will call this new algorithm SMT PCA.

For the OBP1 dataset (Figure 7), the result after starting with the subspace based on the global eigenvectors are better than QLRX when OWS is big (7,9). SMT-PCA gives better results than SMT for any OWS.

For the OBP2 dataset (Figure 8), the result after starting with the subspace based on the global eigenvectors are better than QLRX when OWS is big (7,9). SMT-PCA gives better results from SMT for any OWS (for OWS = 3 SMT and SMT-PCA almost the same).

Examining the number of rotations needed in the SMT and in the SMT-PCA (see Table 3).

TABLE 3: This table shows the time it took to complete the calculation of a dataset.

“Name”	“OWS”	“GWS”	Original SMT number of rotations	SMT after PCA number of rotations	Rotations number ratio
“OP1_T1_S10”	9	1	3845	1851	2.1
“OP1_T1_S33”	9	1	3794	1831	2.1
“OP1_T1_S100”	9	1	3820	1806	2.1
“OP1_T1_S10”	7	1	2781	1508	1.8
“OP1_T1_S33”	7	1	2769	1498	1.8
“OP1_T1_S100”	7	1	2770	1457	1.9
“OP1_T1_S10”	5	1	1965	1177	1.7
“OP1_T1_S33”	5	1	1969	1203	1.6
“OP1_T1_S100”	5	1	1965	1122	1.8
“OP1_T1_S10”	3	1	1195	442	2.7
“OP1_T1_S33”	3	1	1187	407	2.9
“OP1_T1_S100”	3	1	1201	421	2.9
“OP2_T1_S10”	9	1	2135	839	2.5
“OP2_T1_S33”	9	1	2105	840	2.5
“OP2_T1_S100”	9	1	2095	827	2.5
“OP2_T1_S10”	7	1	1480	647	2.3
“OP2_T1_S33”	7	1	1481	641	2.3
“OP2_T1_S100”	7	1	1464	635	2.3
“OP2_T1_S10”	5	1	961	440	2.2
“OP2_T1_S33”	5	1	962	435	2.2
“OP2_T1_S100”	5	1	960	431	2.2
“OP2_T1_S10”	3	1	556	203	2.7
“OP2_T1_S33”	3	1	555	205	2.7
“OP2_T1_S100”	3	1	552	206	2.7

The smaller the number of rotations, the less time needed for the calculation.

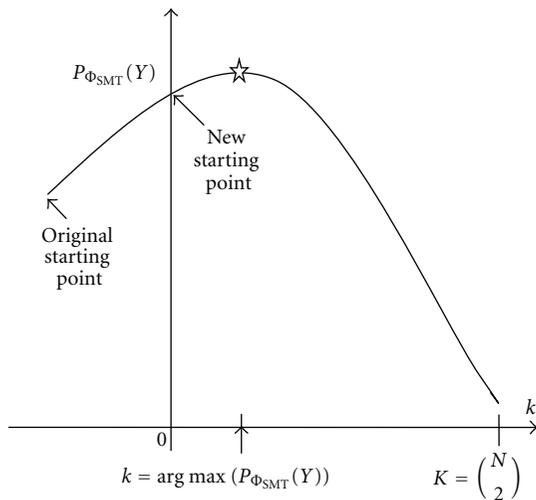


FIGURE 6: When starting from PCA subspace, we will start from a closer point to the maximum so we need fewer rotations; the delta in k between the original location to the current one is the rotations done by transforming to the PCA subspace.

6. Conclusions

As a preliminary to our conclusions, please note that when we discuss using a small or large number of pixels, that in all

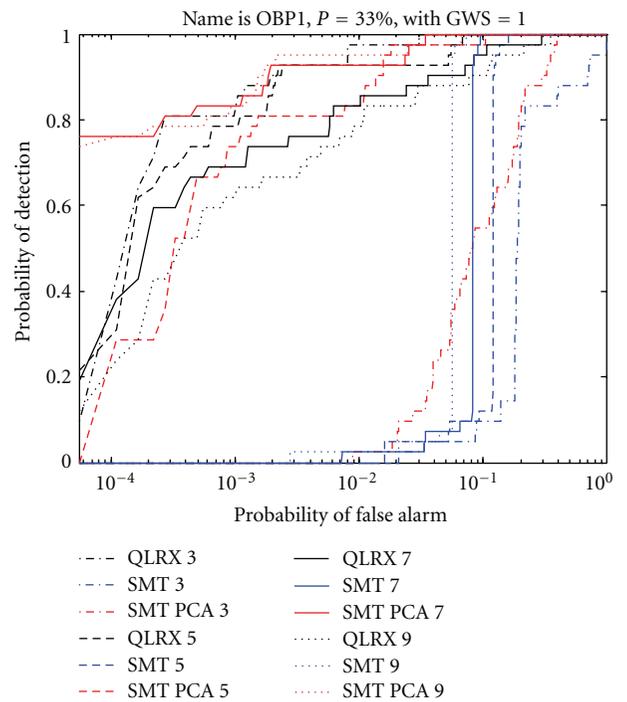


FIGURE 7: OBPI results with OWS given by the number and the legend.

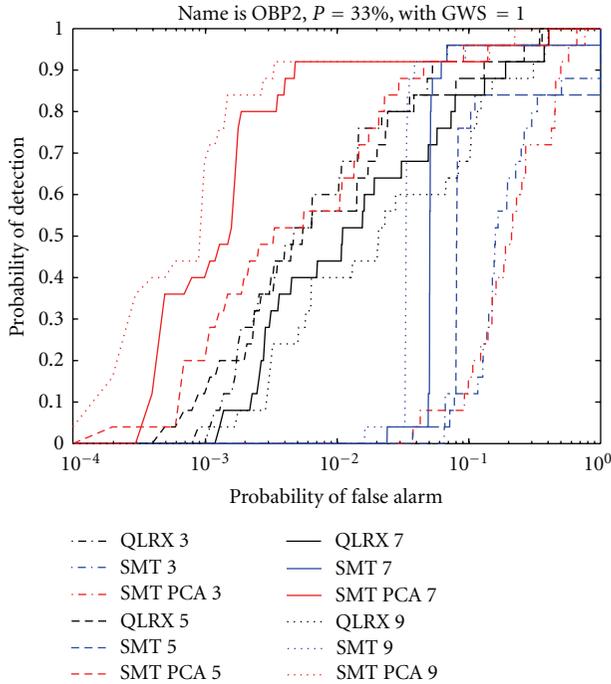


FIGURE 8: Similar to Figure 7 for the HAR dataset.

cases the number of pixels used is less than the number of spectral bands.

Two methods in this paper have been considered for dealing with possibly singular covariance matrices. In the first (QLRX), we use global eigenvectors and local eigenvalues as an approximation of the inverse covariance matrix. In the second (SMT), we use an iterative process to slowly “twist” our axes to come closer to those determined by the data.

In our two datasets, we found that if a small area was used for estimating the background, the QLRX algorithm was superior. For large areas of background, QLRX remains superior, although SMT greatly improves as follows:

$$\frac{\text{Number of pixels}}{\text{Number of bands}} = \frac{p}{n} \in (0, 1), \quad (16)$$

- (i) the calculation time of QLRX is much smaller (two orders of magnitude) than both SMT and SMT-PCA,
- (ii) the calculation time of SMT PCA is less than the calculation time of the original SMT by about a factor of two,
- (iii) SMT-PCA and QLRX performance are better than those of SMT for any number of pixels,
- (iv) for a small number of pixels ($p/n \leq 0.1$), the QLRX performance is better than that of SMT-PCA,
- (v) for a large number of pixels ($0.2 \leq p/n < 0.1$), the performance of SMT-PCA is better than that of QLRX.

Acknowledgments

The authors gratefully recognize the partial support for this work from the Paul Ivanier Center for Robotics Research and Production Management, Beer-Sheva, Israel. The test images are part of the Oberpfaffenhofen HyMAP scene, collected in 2004, during an airborne campaign sponsored by the Belgian Science Policy Office (BelSPO). Flights were operated by the German Aerospace Center (DLR).

References

- [1] C. E. Cafer, M. S. Stefanou, E. D. Nielsen, A. P. Rizzuto, O. Raviv, and S. R. Rotman, “Analysis of false alarm distributions in the development and evaluation of hyperspectral point target detection algorithms,” *Optical Engineering*, vol. 46, no. 7, Article ID 076402, 2007.
- [2] C. E. Cafer, J. Silverman, O. Orthal, D. Antonelli, Y. Sharoni, and S. R. Rotman, “Improved covariance matrices for point target detection in hyperspectral data,” *Optical Engineering*, vol. 47, no. 7, Article ID 076402, 2008.
- [3] J. Theiler, G. Cao, L. R. Bachega, and C. A. Bouman, “Sparse matrix transform for hyperspectral image processing,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 424–437, 2011.
- [4] D. Borghys and C. Perneel, “Study of the influence of pre-processing on local statistics-based anomaly detector results,” in *Proceedings of the Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS '10)*, pp. 1–4, June 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

